

Alexandra Emmons, PhD
Bioinformatics Training and Education Program (BTEP)
ncibtep@nih.gov

An Introduction to QIIME2

Cancer and the microbiome

What is QIIME2?

QIIME2 Core Installation

Tools for reproducibility and visualization

Showcase of core methods and visualizations

Community Support

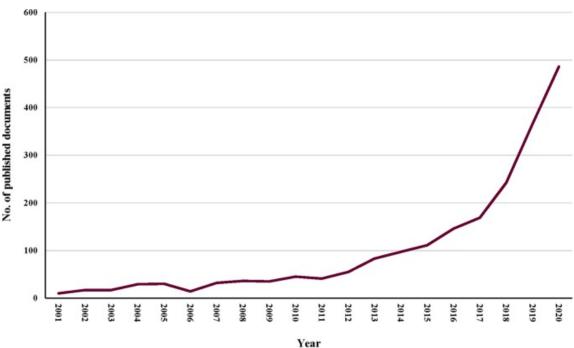
Workshops and tutorials

The future of QIIME2

Trends in Microbiome Research

Microbiome – a collection or community of microorganisms in a defined environment





Growth trends of publications on the microbiota and cancer from 2001 to 2020, Zyoud et al. 2022, *Journal of Translational Medicine*

The role of the microbiome in cancer

The microbiome, cancer, and cancer therapy

Beth A. Helmink, M. A. Wadud Khan, Amanda Hermann, Vancheswaran Gopalakrishnan & Jennifer A. Wargo ⊡

Nature Medicine 25, 377–388 (2019) | Cite this article 52k Accesses | 321 Citations | 240 Altmetric | Metrics

Liver/biliary tract Breast Exposure to bacteria Affects levels of and/or metabolites via circulating estrogens Alters balances in portal venous system energy metabolism Production of secondary bile acids may alter immune Colon function (NKT cells) Chronic inflammatory influencing tumor state induced by toxins growth (primary and or alterations in metastatic) signaling pathways DNA damage and Potentiation of hepatotoxicity induced inflammatory state by by secondary bile acids breakdown of normal and other mechanisms mucosal barrier Induction of nonalcoholic Production of toxins that steatohepatitis and are directly genotoxic other conditions that Increases local predispose to cirrhosis production of reactive and subsequent cancer oxygen species formation Induction of oncogenic transcriptional activity potentially via Stomach β-catenin-Wnt · Induction of a chronic signaling pathways inflammatory state Impairment of Direct genotoxic effects antitumor immune on gastric mucosa function

Fig. 1 | The influence of gut microbiota on cancer development. Shown are various mechanisms through which dysbiosis is proposed to affect tumorigenesis and/or tumor growth across cancer types, including colon, hepatobiliary, gastric and breast. Credit: Debbie Maizels/Springer Nature.

The role of the microbiome in cancer

The microbiome, cancer, and cancer therapy

Beth A. Helmink, M. A. Wadud Khan, Amanda Hermann, Vancheswaran Gopalakrishnan & Jennifer A. Wargo □

<u>Nature Medicine</u> **25**, 377–388 (2019) | <u>Cite this article</u> **52k** Accesses | **321** Citations | **240** Altmetric | <u>Metrics</u>

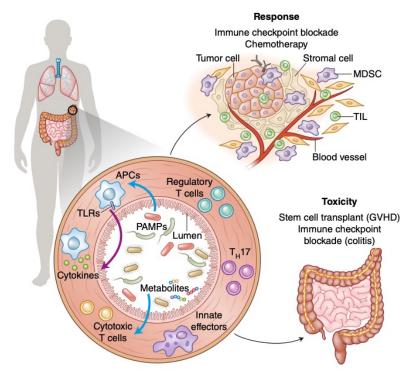
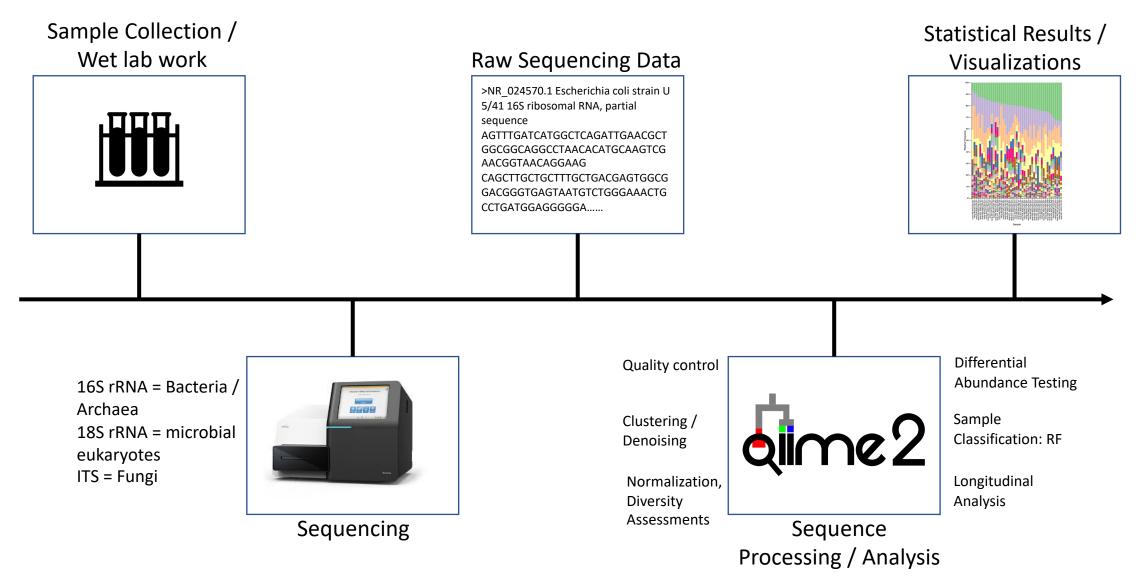


Fig. 2 | The influence of the gut microbiota on different cancer therapies.

Gut microbes can impact both the response to various cancer therapies and associated toxicities, such as colitis and GVHD. The gut microbiota is thought to alter systemic immune function via local changes within the gut mucosa and gut-associated lymphoid tissue. The interaction of PAMPs with APCs and innate effectors via PRRs (TLRs) can help prime an adaptive immune response. Cytokines and microbial metabolites produced locally can act systemically. These combined activities lead to increased antitumor immune function with increased numbers of tumor-infiltrating lymphocytes (TILs) and decreased numbers of myeloid-derived suppressor cells (MDSCs). Credit: Debbie Maizels/Springer Nature.

Basic steps in a microbiome analysis



What is QIIME2?

A powerful, extensible, and decentralized microbiome analysis package with a focus on data and analysis transparency. QIIME 2 enables researchers to start an analysis with raw DNA sequence data and finish with publication-quality figures and statistical results. --- 2016-2021, QIIME 2 development team.



~32,558 citations of QIIME and QIIME2 – Google Scholar, March 2022

molBV reveals immune landscape of bacterial vaginosis and predicts human papillomavirus infection natural history

Mykhaylo Usyk, Nicolas F. Schlecht, Sarah Pickering, LaShanda Williams, Christopher C. Sollecito, Ana Gradissimo, Carolina Porras, Mahboobeh Safaeian, Ligia Pinto, Rolando Herrero, Howard D. Strickler, Shankar Viswanathan, Anne Nucci-Sack, Angela Diaz, Costa Rica HPV Vaccine Trial (CVT) Group & Robert D. Burk

✓

Nature Communications 13, Article number: 233 (2022) | Cite this article



Article | Open Access | Published: 28 September 2021

Gut bacteria identified in colorectal cancer patients promote tumourigenesis via butyrate secretion

Shintaro Okumura, Yusuke Konishi, ... Eiji Hara ☐ + Show authors

Nature Communications 12, Article number: 5674 (2021) | Cite this article

13k Accesses | 9 Citations | 76 Altmetric | Metrics

Gene

Volume 811, 15 February 2022, 146083



Altered cervicovaginal microbiota in premenopausal ovarian cancer patients

Asuka Morikawa ^{a, b}, Ayako Kawabata ^a, Katsuhiko Shirahige ^b, Tetsu Akiyama ^b, Aikou Okamoto ^a, Takashi Sutani ^b $\stackrel{\wedge}{\sim}$

Research | Open Access | Published: 04 January 2022

Urogenital Microbiota: Potentially Important Determinant of PD-L1 Expression in Male Patients with Non-muscle Invasive Bladder Cancer

Chunxiao Chen, Zehai Huang, Pengcheng Huang, Kun Li, Jiarong Zeng, Yuehui Wen, Biao Li, Jie Zhao 🖂 & Peng Wu ⊠

BMC Microbiology 22, Article number: 7 (2022) | Cite this article

578 Accesses | **10** Altmetric | Metrics

Article Open Access Published: 30 November 2021

Oral microbiome associated with lymph node metastasis in oral squamous cell carcinoma

Young-Gyu Eun, Jung-Woo Lee, Seung Woo Kim, Dong-Wook Hyun, Jin-Woo Bae & Young Chan Lee 🖂

Scientific Reports 11, Article number: 23176 (2021) | Cite this article

911 Accesses | 1 Citations | 1 Altmetric | Metrics

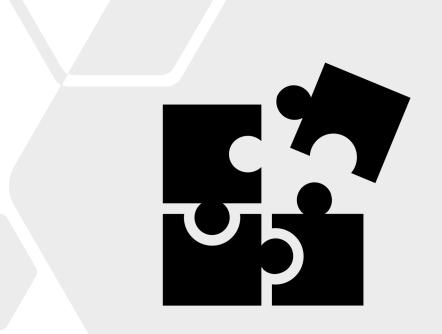
Why use Qiime2?

- Methods evolve with the field
 - Plug-in architecture
- Interactive data exploration
- Easy to use and accessible
 - Open-source
 - Extensive documentation
 - Tutorials and workshops
- Widespread use
 - Qiime2 forum
 - Global network of collaborators

- Provenance tracking
 - Facilitates reproducibility
- Multiple user interfaces
 - Galaxy
 - Q2 Studio prototype
 - Python API
 - Command line
 - R API (IN THE WORKS)

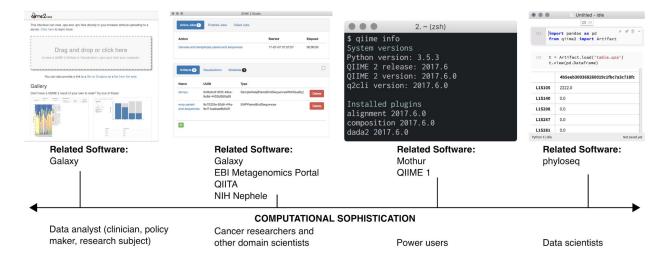
Plug-in Architecture

- Plugins are software packages that can be developed by anyone
- All bioinformatic functionality in QIIME2 defined by plugins
- New plug-ins found in the QIIME2 library



Interfaces

- Q2studio
 - Graphical user interface
- Q2cli
 - Command line interface
 - The classic way to use qiime2
- Artifact API
 - Python 3 application programming interface (API) for QIIME 2
 - Use with Jupyter notebook
 - Recommended for advanced users
- Now on Galaxy
- Potentially coming soon to R
 - For now, qiime2R package



Qiime2 developers, 2021 workshop content

Qiime2 Documentation



Version: 2022.2 ▼

Table of Contents

- Getting started
- What is QIIME 2?
- Core concepts
- Installing QIIME 2
- Jupyter Book Tutorials
- Tutorials
- Interfaces
- Plugins
- Semantic types
- Community
- Data resources
- Supplementary resources
- User Glossary
- Citing QIIME 2

Quick search

Go

QIIME 2 user documentation

This site is the official user documentation for QIIME™ 2, including installation instructions, tutorials, and other impoi information. Visit http://qiime.org for information on QIIME™ 1.

Getting started

Check out the getting started guide to begin using QIIME 2.

Table of contents

- Getting started
- What is QIIME 2?
- Core concepts
 - Data files: QIIME 2 artifacts
 - Data files: visualizations
 - Semantic types
 - Plugins
 - Methods and visualizers
 - Next steps
- Installing QIIME 2
 - Natively installing QIIME 2
 - Installing QIIME 2 using Virtual Machines
 - Recommendations
 - QIIME 2 Core 2022.2 distribution
- Jupyter Book Tutorials
- Tutorials
 - Overview of QIIME 2 Plugin Workflows
 - QIIME 2 for Experienced Microbiome Researchers
 - "Moving Pictures" tutorial

Qiime2 Installation

```
/usr/local/lmod/modulefiles ---
QIIME/1.9.1
                   QIIME/2-2020.8
                                           qiime/2-2019.4
QIIME/2-2018.2
                   QIIME/2-2020.11 (D)
                                           qiime/2-2019.10
QIIME/2-2019.1
                   QIIME/2-2021.4
                                           qiime/2-2020.2
QIIME/2-2019.4
                   qiime/1.9.1
                                           qiime/2-2020.8
                   qiime/2-2018.2
QIIME/2-2019.10
                                           qiime/2-2020.11 (D)
QIIME/2-2020.2
                   qiime/2-2019.1
                                           qiime/2-2021.4
```

q2cli installation on macOS, Linux, and Windows (WSL)

Qiime2 installed on Biowulf

module load QIIME

Getting a Biowulf account:

https://hpc.nih.gov/docs/accounts.html

Qiime 2 Biowulf help page

https://hpc.nih.gov/apps/QIIME.html#:~:text=QIIME2%20on%20Biowulf&text=QIIME%202%20is%20a%20powerful,quality%20figures%20and%20statistical%20results.

Tools for Reproducibility

Qiime2view



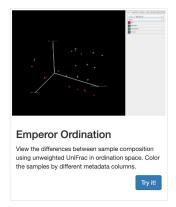
You can also provide a link to a file on Dropbox or a file from the web.

Gallery

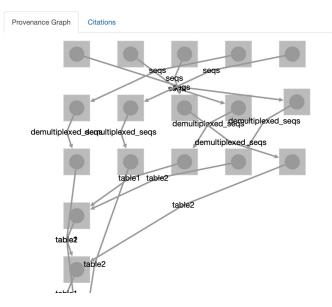
Don't have a QIIME 2 result of your own to view? Try one of these!







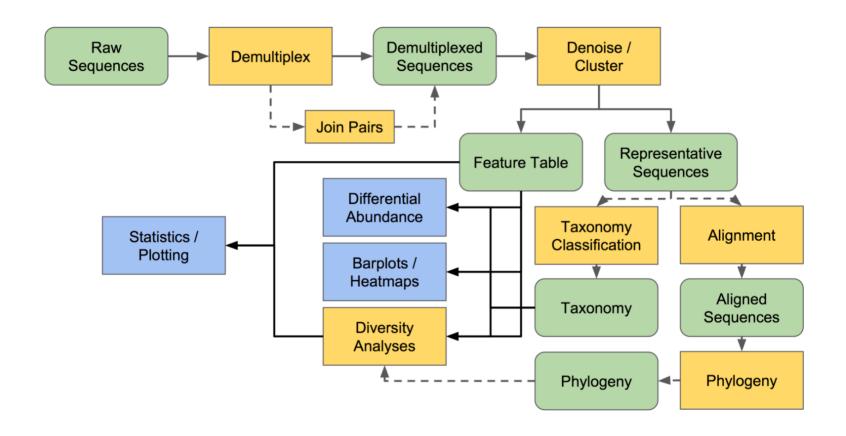
Provenance Tracking



Details

Click on an element of the Provenance Graph to learn more.

Q2 workflows



https://docs.qiime2.org/2022.2/tutorials/over view/#let-s-get-oriented-flowcharts

Core plug-in Overview

https://docs.qiime2.org/2022.2/plugins/available/

Available plugins and methods with q2 installation

- Focus on amplicon sequence processing and analysis
 - Primary targets: 16S rRNA, 18S rRNA, ITS
 - QC analysis, adapter / primer trimming via cutadapt
 - OTU clustering, Sequence Denoising
 - DADA2, Deblur, OTU clustering via vsearch (close reference and open reference)
 - Taxonomic Classification
 - Tree building
 - MAFT and fasttree
 - Raxml
 - SEPP fragment insertion
 - Data analysis
 - RF classification, differential abundance testing, core features, alpha diversity, beta diversity, longitudinal methods.

Data import types

- First step = data import
 - Can import at most stages in the analysis
 - qiime tools import –h
 - Raw data usually some type of fastq files
 - EMPPairedEndSequences
 - EMPSingleEndSequences
 - SampleData[SequencesWithQuality]
 - MultiplexedPairedEndBarcodeInSequence
 - RawSequences
- See the import tutorial for help
 - https://docs.qiime2.org/2022.2/tutorials/ importing/

qiime tools import --show-importable-types

```
Bowtie2Index
DeblurStats
DistanceMatrix
[EMPPairedEndSequences
EMPSingleEndSequences
ErrorCorrectionDetails
FeatureData[AlignedProteinSequence]
FeatureData[AlignedRNASequence]
FeatureData[AlignedSequence]
FeatureData[Differential]
FeatureData[Importance]
FeatureData[PairedEndRNASequence]
FeatureData[PairedEndSequence]
FeatureData[ProteinSequence]
FeatureData[RNASequence]
FeatureData[SILVATaxidMap]
FeatureData[SILVATaxonomy]
FeatureData[Sequence]
FeatureData[Taxonomy]
FeatureTable[Balance]
FeatureTable[Composition]
FeatureTable[Design]
FeatureTable[Frequency]
FeatureTable[PercentileNormalized]
FeatureTable[PresenceAbsence]
FeatureTable[RelativeFrequency]
Hierarchy
MultiplexedPairedEndBarcodeInSequence
MultiplexedSingleEndBarcodeInSequence
PCoAResults
Phylogeny[Rooted]
Phylogeny[Unrooted]
Placements
ProcrustesStatistics
QualityFilterStats
RawSequences
SampleData[AlphaDiversity]
SampleData[BooleanSeries]
SampleData[ClassifierPredictions]
SampleData[DADA2Stats]
SampleData[FirstDifferences]
SampleData[JoinedSequencesWithQuality]
SampleData[PairedEndSequencesWithQuality]
SampleData[Probabilities]
SampleData[RegressorPredictions]
SampleData[SequencesWithQuality]
SampleData[Sequences]
SampleEstimator[Classifier]
SampleEstimator[Regressor]
SeppReferenceDatabase
TaxonomicClassifier
UchimeStats
```

Demultiplex methods & Quality Control

Demultiplex: Associating sequences with samples

- May be done by the sequencing core
- Methods available
 - q2-demux
 - Emp-paired
 - Emp-single
 - q2-cutadapt
 - Demux-paired
 - Demux-single

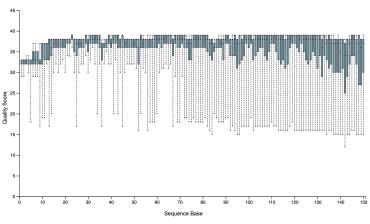
Quality Control

- Adapter, primer trimming
 - Q2-cutadapt
- Quality score filtering
 - q2-quality-filter
- Removing human contamination
 - Q2-exclude-seqs
 - Q2-filter-reads
- Community evaluation (expected vs observed)
 - Q2- quality-control

erview Interactive Quality Plot

Click and drag on plot to zoom in. Double click to zoom back out to full size. Hover ov

Forward Reads



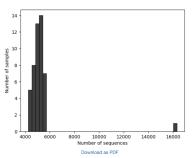
These plots were generated using a random sampling of 10000 out of 253874 sequences without replacement. The minimum sequence length identified during subsampling was 150 bases. Outlier quality scores are not shown in box plots for clarity.

Overview Interactive Quality Plot

Demultiplexed sequence counts summary

	forward reads
Minimum	4237
Median	5101.5
Mean	5289.041667
Maximum	16327
Total	253874

Forward Reads Frequency Histogram

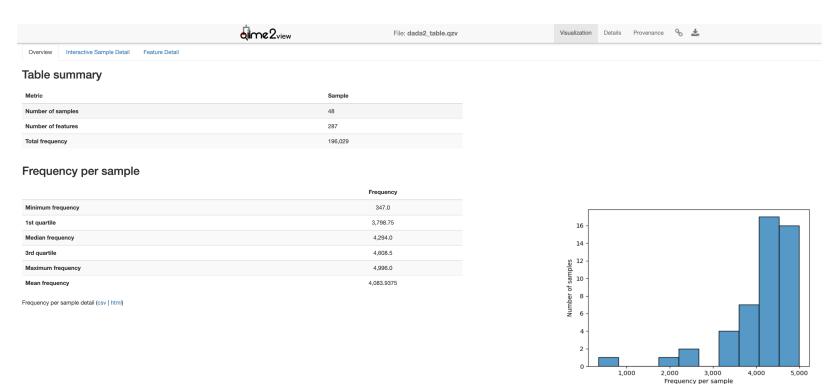


Feature Table generation

- OTU clustering
 - vsearch
- Denoising
 - Deblur
 - DADA2

Outputs: Two key files

- 1. Feature table
- 2. Representative Sequences



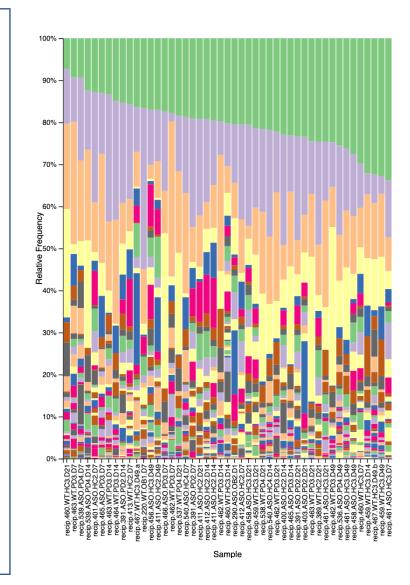
Other feature table methods:

https://docs.qiime2.org/2022.2/plugins/available/feature-table/

Taxonomic Classification & Phylogenetic Tree

Taxonomy

- classify-consensusblast
 - BLAST+ consensus taxonomy classifier
- classify-consensusvsearch
 - VSEARCH-based consensus taxonomy classifier
- classify-sklearn
 - Pre-fitted sklearnbased taxonomy classifier



Tree

- Fragment insertion –
 SEPP
- qiime phylogeny
 - <u>align-to-tree-mafft-</u> fasttree
 - <u>align-to-tree-mafft-</u> iqtree
 - <u>align-to-tree-mafft-</u> raxml

NO CURRENT WAY TO VIEW TREES USING QIIME2 PLATFORM

Community analysis (visualizations / statistics)

qiime diversity core-metricsphylogenetic

Emperor plots

Alpha and beta group significance testing

• output

Differential abundance testing

ANCOM

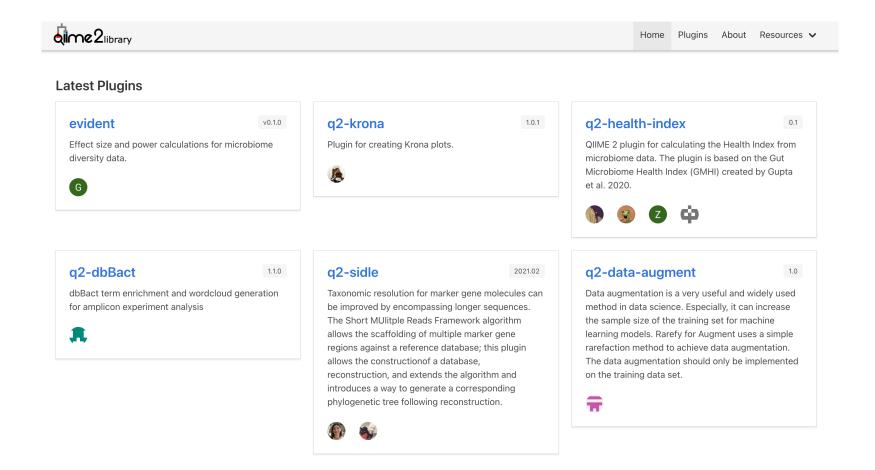
Longitudinal testing

Distance based

Data Resources

- Taxonomy Classifiers
- Weighted Taxonomic Classifiers
- Q2 formatted reference marker gene databases
- SEPP reference databases

QIIME2 Plugin Library



Future Features

- Funding through NCI
 - To increase the accessibility of the QIIME 2 platform and cancer microbiome data to the cancer research community
 - Galaxy
 - R API
 - Cancer microbiome data resource via QIITA
 - cancer-focused QIIME 2 user workshop series
 - Collaboration with NCI Cloud Resources teams to integrate QIIME 2
 - Work directly with data from GATK and TCGA
 - Make QIIME 2 a microbiome mutli-omics bioinformatics platform
 - Facilitate reproducible research

<u>Informatics Technology for Cancer Research</u> program **Project 1U24CA248454-01**;

https://reporter.nih.gov/project-details/9951750

Future – omics tools

- Integrating multi-omics data
 - Shotgun surveys of phage and other viruses
 - Functional potential via shotgun metagenomics
 - Functional activity
 - Metatransciptome
 - Metaproteome
 - metabolome

Provenance Replay

- The ability to generate executable workflows from an existing artifact or visualization's data provenance

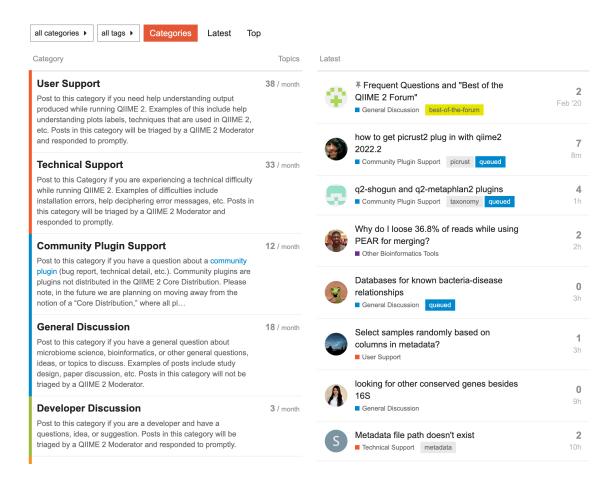
Other features:

- ability to obtain provenance info in a publication ready format
- Automated workflow bug detection from data provenance
- Plugins to help users submit data to repositories

Qiime2 forum

A vast network of collaborators and users

Help in just a short period of time



Current Learning Opportunities through Qiime2



Interested in hosting or attending an official QIIME 2 Workshop? Contact us.

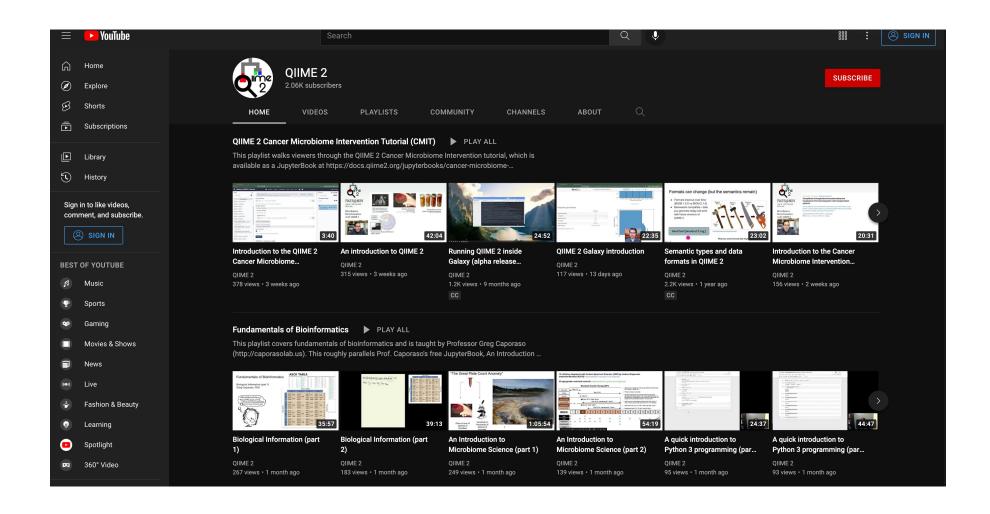
Upcoming Workshops

No upcoming workshops

Past Workshops

Title	Location	Workshop Dates
BIOF 089 - Microbiome Bioinformatics with QIIME 2	Online (via FAES at the National Institutes of Health)	Jan. 31, 2022 - Feb. 4, 2022
Microbiome Bioinformatics with QIIME 2: Free online workshop	Online	Jan. 26, 2021 - June 1, 2021
Microbiome Bioinformatics with QIIME 2	Online (via FAES at the National Institutes of Health)	Jan. 4, 2021 - Jan. 8, 2021
An Introduction to QIIME 2	Online (via World of Microbiome)	Nov. 4, 2020 - Nov. 4, 2020
CZI-CABANA Microbiome Bioinformatics with QIIME 2 Workshop (sold out)	Online	Oct. 5, 2020 - Oct. 9, 2020
Introduction to microbiome study design and analysis	Puerto Rico	Aug. 1, 2020 - Aug. 1, 2020
Microbiome Bioinformatics with QIIME 2	Bethesda, Maryland	Jan. 8, 2020 - Jan. 10, 2020
Microbiome Bioinformatics with QIIME 2 Workshop	Fort Collins, Colorado (USA)	Nov. 25, 2019 - Nov. 26, 2019
Microbiome Bioinformatics with QIIME 2 Workshop (not open to the public)	University of Wyoming	Oct. 24, 2019 - Oct. 25, 2019
Microbiome Bioinformatics with QIIME 2	Bangkok, Thailand	Sept. 11, 2019 - Sept. 12, 2019
QIIME 2 @ One Health Summer School	University of Bern, Switzerland	Aug. 14, 2019 - Aug. 14, 2019

QIIME2 Youtube Channel



New Cancer Microbiome Tutorial using Galaxy!



QIIME 2 Cancer Microbiome Intervention Tutorial

Q Search this book...

QIIME 2 Cancer Microbiome Intervention Tutorial

GETTING STARTED WITH QIIME 2

Getting started with using QIIME 2









QIIME 2 Cancer Microbiome Intervention Tutorial

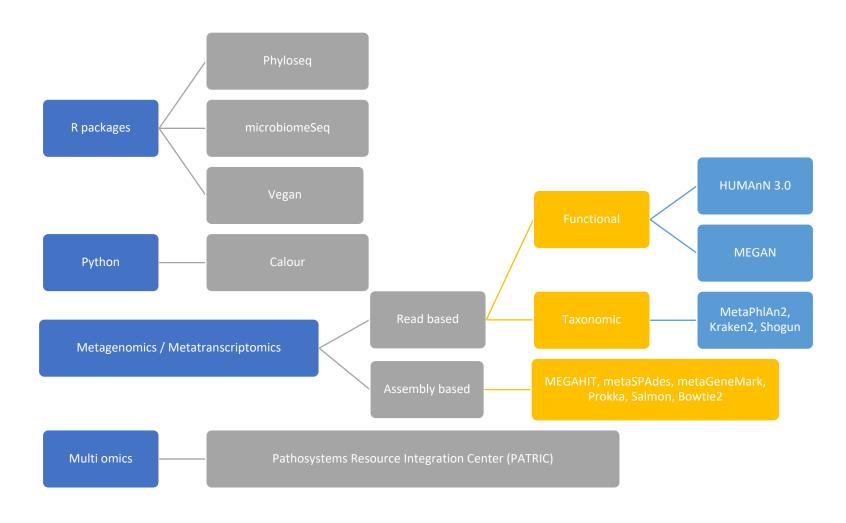
This tutorial was originally developed for the 2022 NIH Foundation for Advanced Education in the Sciences Microbiome Bioinformatics with QIIME 2 workshop, taught online January 31 - February 4, 2022. It is now being publicly released with corresponding video content on the QIIME 2 YouTube channel.

This tutorial is intended to be read like a book, from beginning to end.

This tutorial and the corresponding videos present many new features of QIIME 2 and our education and documentation ecosystem, including:

- The new QIIME 2 Galaxy-based graphical user interface;
- · New tutorial data;
- Our new QIIME 2 documentation system, which illustrates how to use QIIME 2 with different user
 interfaces, including Galaxy, the QIIME 2 command line interface, and the QIIME 2 Python 3 API. You can
 choose to use whichever of these interfaces you're most comfortable with, or user different interfaces for
 different steps;
- Our first Jupyter Book-based QIIME 2 tutorial.

Going beyond qiime2



Contact me at ncibtep@nih.gov for help regarding this resource.